ABSTRACT

A method for pre-processing speech, in particular for recognizing speech, including receiving a speech signal, separating a spectrum of said speech signal into a number of predetermined frequency sub-bands, analyzing said speech signal within each of said frequency sub-bands, generating respective band-dependent acoustic feature data for each of said respective frequency sub-bands, deriving band-dependent likelihoods for occurrences of speech elements or within said speech signal based on said band-dependent acoustic feature data, analyzing said speech signal within said spectrum, generating full-band acoustic feature data, which are at least in part representative for said speech signal with respect to said spectrum, deriving a full-band likelihood for occurrences of speech elements or of sequences thereof within said speech signal based on said full-band acoustic feature data, deriving an overall likelihood for occurrences of speech elements within said speech signal based on said band-dependent likelihoods and said full-band likelihood.